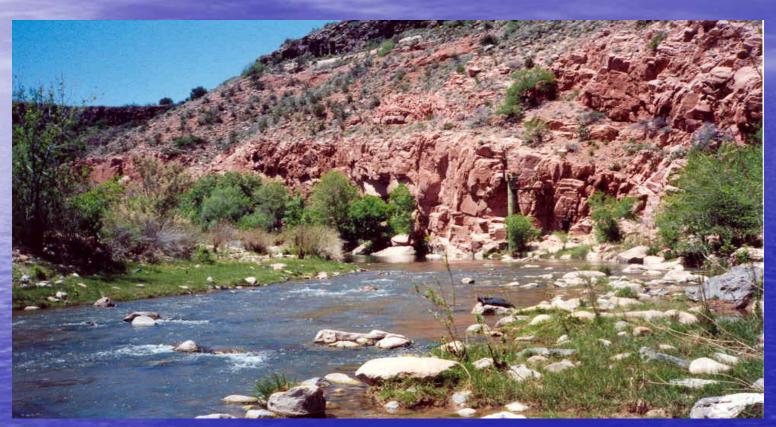
### Impaired Water Identification Rule

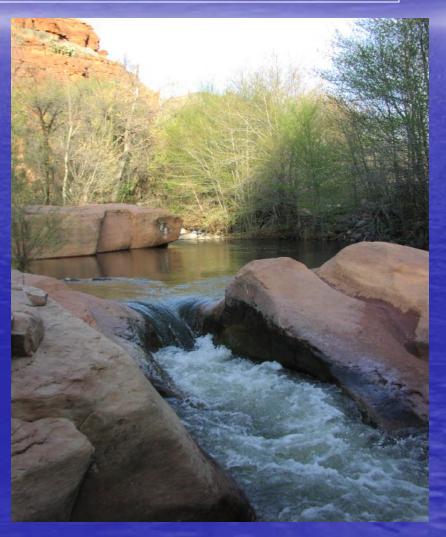


Stakeholder Meeting #2
August 9, 2004

# ADEQ Mission: To protect public health and the environment

### Goal for this process:

To adopt a more scientifically sound rule for identification of impaired waters that is protective of public health and the environment.



### What we need to accomplish...

- Binomial Approach
- Assessment of chronic Aquatic and Wildlife standards
- Assessment of Escherichia coli
- Planning List
- Miscellaneous corrections

Draft rule for review

Publish rule for review

Meetings to discuss rule

Publish notice of proposed rulemaking

**Oral Proceedings** 

File Notice with GRRC

File Rule with Sec. of State

Begin application of new rule!

September 2004

October 2004

October 2004

November 2004

January 2005

February 2005

March 2005

May 2005

## Meeting Topics

### Purpose of today's meeting:

To address comments regarding the binomial approach and introduce proposed changes. Comments to be addressed:

- Minimum sample size
- Representative samples
- Statistical confidence

#### Next meetings:

Chronic standards and *E. coli* standards assessment Planning List and miscellaneous Draft rule

Week of Aug. 23rd

Week of Sept. 13th Week of Sept. 27th

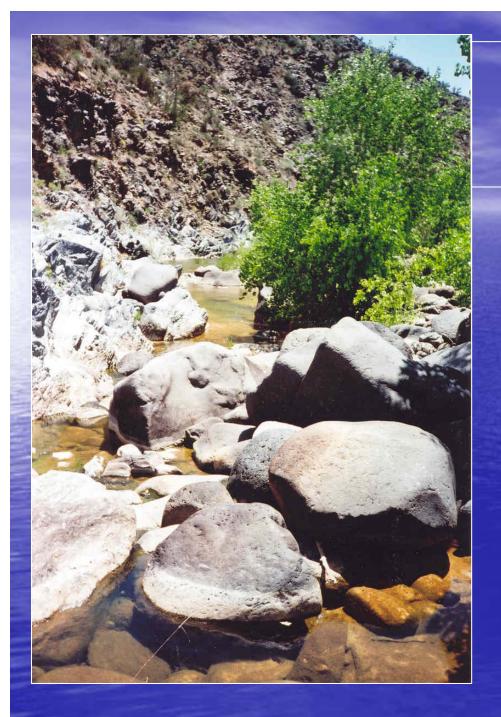
## Agenda

- Assessments 101Melanie Diroll
- Response to Comments:
   Minimum sample size and representative samples
   Melanie Diroll
- Response to Comments:Statistical ConfidenceDiana Marsh
- Proposed Changes: Modified Binomial Approach
  - Diana Marsh
- DiscussionAll

### Clean Water Act Requirements



- CWA §305(b) requires assessment of all surface waters every two years using "all readily available data"
- CWA §303(d) requires list of impaired waters every two years



### **EPA** Oversight

- Assessment and list of impaired waters submitted to EPA Region IX as an Integrated Report
- EPA has maintained authority to approve or disapprove 303(d) list (CFR §130.7)

### What is the goal of assessments?

- To provide an overall picture of the state's ambient, or in-stream, water quality.
- We are not assessing
  - Effluent water at end-of-pipe
  - Storm water run-off before it reaches stream

### Where is the data collected?

- We have about eight ambient monitoring staff to cover the state. Focus is on:
  - Perennial streams and lakes
  - Broad geographic coverage
  - Impacts from nonpoint source pollution
- Seek out data from other sources:
  - Contract with USGS to monitor "non-wadeable" streams
  - Permitted facilities' in-stream data
  - Other agencies (USFWS, USFS, NPS)
  - Universities
  - Citizen volunteers

### What are we assessing?

- Need to know if surface waters are attaining their "designated uses:"
  - Aquatic and Wildlife
  - Fish Consumption
  - Domestic Water Source
  - Recreation
  - Agriculture
- The units of assessment are stream reaches and lakes
- We measure attainment using numeric and narrative water quality standards
- Waterbody can be:
  - Attaining
  - Inconclusive
  - Impaired (includes Not attaining)

### How do we assess?

- ADEQ has an obligation to assess "all" waters every two years
- EPA, in conjunction with state work groups, developed a Consolidated Assessment and Listing Methodology (CALM)
- Given a limited number of samples:
  - What is sufficient evidence to indicate attainment of a des. use?
  - What is sufficient evidence to indicate impairment of <u>at least one</u> des. uses?

### Attaining Assessments

- Set of core parameters
  - What pollutants we expect to find in Arizona
  - What parameters are indicators of stream health
- We need at least 3 samples seasonally distributed to call something "attaining all uses," provided no exceedances are found

### Impaired Assessments

- Impairment: repeated or persistent violations of WQ standards which have a negative impact on one or more designated uses
- Statute adopted in 2000 required that we establish in rule:
  - Credible data requirements for listing
  - Minimum sample size requirements for listing
  - 303(d) listing and delisting methods
- Impaired Water Identification Rule was adopted in 2002

### Methods to determine impairment

- Statistical method Binomial Approach
  - >10% exceedance rate at given confidence level (90%)
- Method for "acutely toxic" parameters
  - >1 exceedance in 3 years
- Method for statistically-based standards
  - >1 exceedance after statistical calculation is applied

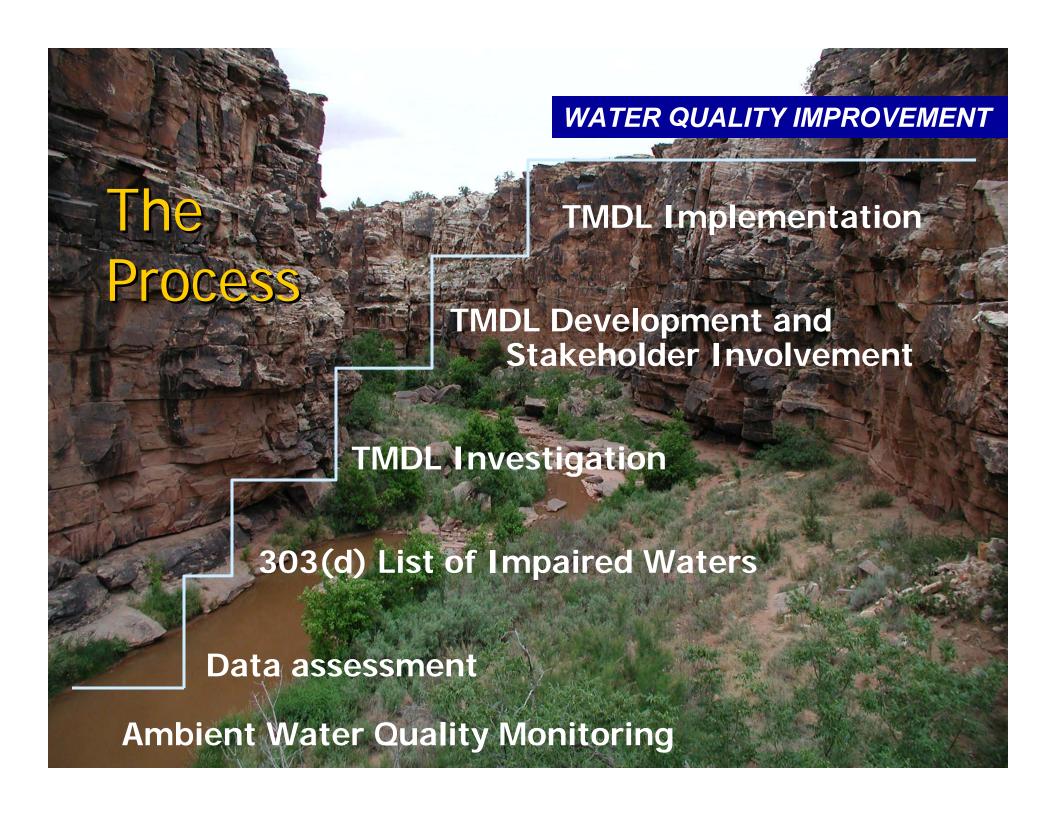
# Impairment: What happens next?

### TMDL Investigation

- Placement on list does not require any changes to occur in the watershed
- Simply requires ADEQ to conduct further monitoring
- Intensive water quality study identify and quantify sources
- Existing permits do not change
- Waterbody can be delisted if no longer impaired or listing was in error

### TMDL Development and Implementation

- Stakeholder involvement is essential.
- Landowners, permittees, and communities work with ADEQ to develop a plan to reduce pollutant loadings
- Grant monies are available for nonpoint source pollution



### Stakeholder Concerns

- Adherence to credible data requirements in statute:
  - Adequate number of samples
  - Representative samples
  - Appropriate QA/QC procedures
  - Valid sampling and analytical methods
- Comments:
  - How many samples should it take to call a waterbody impaired?
  - How often should samples be collected?
  - Over what length of time?
  - Over what area?

# Number of Samples: How Many?

Comment:

Need many samples to accurately represent an entire stream reach or lake and call it impaired.

Statistical tests have good power for detecting exceedances if they are based on 30 or more samples. Smaller sample sizes are prone to yield erroneous <a href="mailto:attainment">attainment</a> decisions because they have a <a href="mailto:low">low</a> probability for detecting exceedances.... (CALM)

This says we need many samples to assess "attaining." Does the same logic apply to assessing impairment?

# Number of Samples: How Many? Assessing Health vs. Assessing Illness

- How does a doctor show you are HEALTHY?
  - One indicator: body temperature
  - Go for 3 checkups in 5 yrs., do not have a fever
  - Is this sufficient evidence to say you never had a fever?
  - NO

- How do we show a waterbody is HEALTHY?
  - One WQ standard: lead

- Go for 3 checkups in 5 yrs., no violations
- Is this sufficient evidence to say you never had a violation?
- NO

# Number of Samples: How Many? Assessing Health vs. Assessing Illness

- How does a doctor show you are ILL?
  - One indicator: body temp.
  - Go for 3 checkups in 5 yrs.
     and you do have a fever each time
  - Is this sufficient evidence to say you were ill?
  - YES

- How do we show WQ violations ARE occurring?
  - One WQ standard: lead
  - Visit 3 times, find 3 violations
  - Is this sufficient evidence to say you have a WQ problem?

YES

# Number of Samples: How Many? Assessing Health vs. Assessing Illness

#### What the CALM Guidance Means

- Small sample sizes have low probability of detecting exceedances...
  - If you take only 3 samples, you may have missed a problem
- If you have 3 bad visits, chances are good those weren't the only 3 times
  - Chances of finding a problem in 3 out of 3 RANDOM visits is VERY SMALL
- CALM is <u>not</u> saying if you find a problem 3, 5, 10 times, that you must wait until 30 visits have taken place before assessing impairment.



How can we assess anything as attaining without constant sampling?

ADEQ Ambient Monitoring vs.

Monitoring below a discharge



- CALM Guidance recommends an assessment of impaired based on 10% exceedance rate, no minimum sample size.
- EPA TSD for Toxics Control recommends minimum of 4 to 6 daily grab samples per month.
- Consider context of guidance
  - TSD "provides guidance for assessing and regulating the discharge of toxic substances to the waters of the U.S."

### Routine Monitoring vs. Monitoring below a Discharge

- Assessing health when risk factors are minimal
  - A healthy person with minimal risk factors
  - Good check-up every few years
  - Can we be reasonably assured you're healthy? YES

- Assessing water quality when risk factors are minimal
  - ADEQ routine monitoring
  - Visit a stream 3 times during multiple seasons, no exceedances
  - Can we be reasonably assured it is healthy? YES

### Routine Monitoring vs. Monitoring below a Discharge

- Assessing health when risk factors are present
  - What if you're over 65 and have a family history of heart problems?
  - Should you go in for a checkup every few years?
  - NO, frequent visits and more tests are warranted

- Assessing water quality when risk factors are present
  - Monitoring below a discharge
  - Visit 3 times?
  - NO, frequent visits and more tests are warranted

- Frequent monitoring is necessary in these areas to show that violations are NOT OCCURRING
- It does not take so many samples to show that violations ARE OCCURRING.

# Number of Sample Sites: How Many?

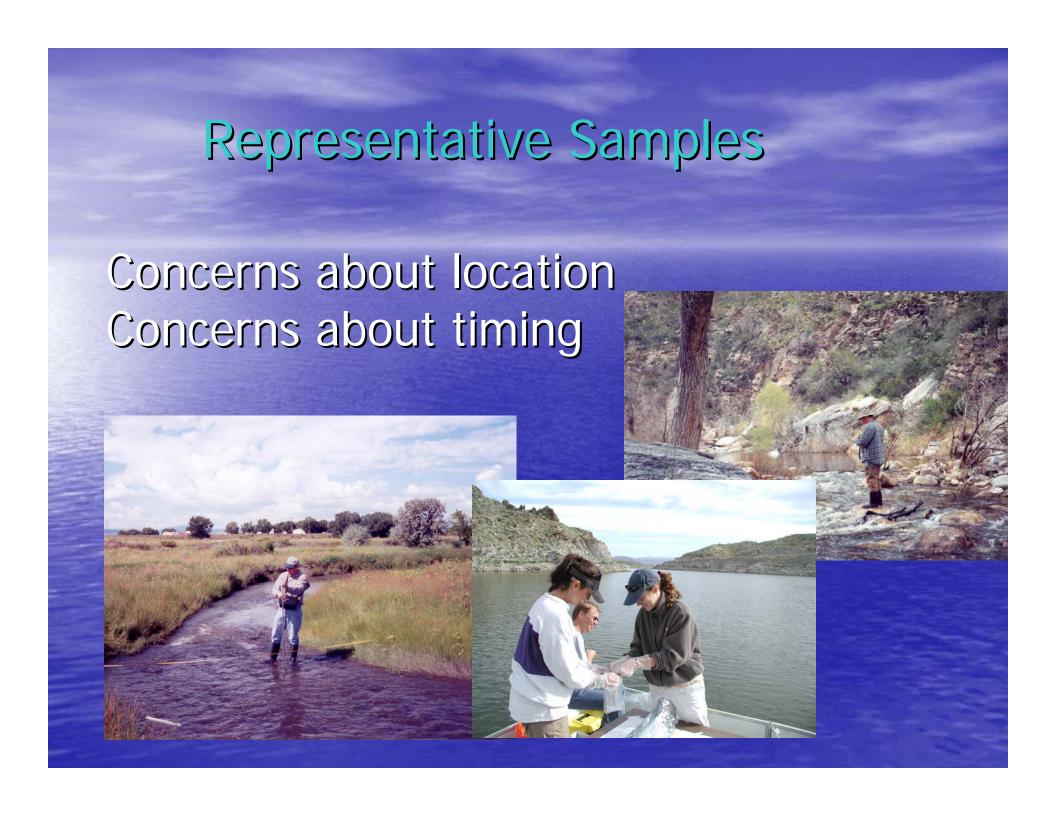
#### Comment:

The minimum sample size of 20 is needed to represent the entire stream reach or lake.

#### Comment:

Consider different minimum sample sizes based on size of reach or lake.

Both valid points, but...
 ADEQ must determine whether ANY PORTION of the waterbody is impaired.



### Representative Samples: Location Concerns

- Comment:
  - What if samples were taken in a "bad spot?"
    - Simple answer is that standards apply to the entire waterbody
- But there are exceptions:
  - Where you have stream conditions that would naturally not meet standards, and no potential pollution source is present.
    - Dissolved oxygen violations in stagnant pools
    - Dissolved oxygen violations at a ground water upwelling
- Comment was made that rule must therefore require a high number of samples for listings (20)?
  - Does this really address the problem?
  - Not really. If we get five D.O. violations, further sampling up to 20 does not change the assessment

### Representative Samples: Location Concerns

- These errors are managed by:
  - Step 1: Check sample plans sites must be representative.
  - Step 2: Weight-of-evidence check individual D.O. violations flow, DB comments, check for possible sources, talk to samplers.
  - Step 3: Release draft 303(d) List opportunity to review and submit documentation that samples or sample sites are not "representative."
- If the waterbody still makes it on the list:
   A delisting can be made during TMDL investigation.
   Sometimes it takes further study to determine whether the impairment is naturally-caused.

### Representative Samples: Timing Concerns

- Comment:
  - What if the samples were taken at a "bad time?"
  - Again, simple answer is standards apply at all times
- But what about temporary or one-time exceedances of standards?
  - Spills that have been remediated
  - Actions have been taken that have improved water quality since samples were collected

## Representative Samples: Timing Concerns

- Current IWIR provides "outs:"
  - A listing will not be made if the surface water is expected to attain its designated use by the next assessment (R18-11-604(D)(2)(h)).
  - The weight-of-evidence approach newer measurements are weighted heavier than older (R18-11-605-(B)(1)(c)).
- If the waterbody still makes it onto the 303(d) list, the rule then provides for delisting when the data used for listing is superseded by more recent data showing standards are attained (R18-11-605(E)(2)(a)(ii)).
- We want to make listings only when problem is recurring or persistent

### Determination of Impairment

## In summary, It is ADEQ's responsibility to:

- Determine whether <u>ANY PORTION</u> of a reach or lake is impaired, and
- Determine whether it is impaired AT ANY TIME (provided the problem is persistent or recurring).

### Determination of Impairment

- If at one site on a 10 mile stream reach, we find repeated WQ violations...
  - Does this constitute impairment? Yes.
  - The TMDL study will confirm the source and delineate the extent of the impairment.
- If a stream reach is attaining its uses during most of the year, but repeatedly violates standards during times of heavy recreational use...
  - Does this constitute impairment? Yes.
  - TMDL study and implementation plan will address these critical conditions.

## How Determine Impairment?

	Parameters and Uses	Assessment Method
Acutely Toxic	Aquatic and Wildlife acute standards, bacteria, nitrate	> 1 exceed in last 3- years of monitoring
Statistically-based Standard	Bacterial geometric mean, nutrient annual mean, etc	> 1 exceed in assessment period (5 years)
All Others	Conventional parameters (DO, pH), human health standards, agricultural standards)	Binomial – Statistically-based assessment

## Binomial Approach

- EPA assessment guidance -- impaired if >10% of samples exceed standards.
  - 1 of 10 samples is 10% not impaired
  - 2 of 10 samples is >10% impaired
- The binomial
  - Use the statistics of a binomial distribution curve to determine the number of samples and exceedances at an acceptable confidence level.
  - By establishing a confidence level, we correct for potential sampling error.
  - For example, at a 95% confidence level, >10% exceedances becomes 3 of 10 samples - impaired

## Binomial Approach

- Existing Binomial
  - 20 samples and 5 exceedances minimum

#### Concerns:

- Is Arizona's binomial sufficiently protective?
   Why are we not listing on 5 of 5 exceedances?
   Should the 20 sample minimum be reduced?
- Is it statistically sound?What is the statistical basis of the minimums?
- What are other states and regions doing?
   How does our binomial compare?

### Show us the Statistics

- Our binomial was modeled after the binomial developed by Florida's Dept of Environmental Protection.
- Florida State University (FSU) Dept of Statistics provided the statistical basis for listing and delisting criteria. Their report's conclusions were:
  - for 95% confidence that >10% exceedance has occurred:
    - a. The minimum number of samples is 10.
    - b. The minimum number exceedances is 3.

## Minimum Sample Size

- ADEQ can justify using 10 instead of 20 samples now because:
  - Data must meet credible data requirements.
  - IWIR has a provision to not list if exceedances represent temporary conditions, such as a spill
  - Other provisions in IWIR, statute, and surface water standards allow ADEQ to not list if impairment due to natural conditions, other actions will result in standards being met, etc.

## Minimum Sample Size

Minimum Sample Sizes and Exceedances for Listings by Other States
When Using a Binomial (b) or Straight >10% (s)

States	AL (s)	KS (b)	NH (b)	NM (s)	NC (s)	SC (s)	TX (b)	VA (s)	WY (s)
Minimum Samples	5	12	10	7	10	2	10	12	2
Minimum Exceed	5	3	2	2	2	2	3	2	2

### Confidence Level

- ADEQ proposes to raise the confidence level from 90% confidence to 95% confidence.
  - This would conform to the confidence level recommended in the FSU report and statistically increase the reliability of assessments.
  - The change would require more samples exceed standards before a listing is made.
  - This change would also increase minimum sample size to support delistings (once impaired).

### Binomial Tables

#### Five tables:

- Table 1 Existing impaired water binomial table (from IWIR)
- Table 2 Revised impaired water binomial table
- Table 3 -- Binomial basis for determining use <u>attainment</u>
- Table 4 Combines Table 2 and 3, to indicate when "inconclusive"
- Table 5 Binomial basis for determining <u>delisting</u> after a surface water is listed as impaired.

### Binomial Tables

- Tables 2 Revised binomial table for determining <u>impaired</u> <u>waters.</u>
- Created using the CRITBINOM function in Excel
  - CRITBINOM (samples, 0.105, 0.95),
  - Calculation provides the minimum number of exceedances for a given sample size, when 95% confidence that >10% exceedance is occurring.
- Table 2 differs from existing binomial (Table 1):
  - 95% instead of 90% confidence level
  - Calculated using 10.5% for >10% rather than adding an exceedance to the value calculated at 10%
  - Used a minimum of 10 samples rather than 20 samples.

### Binomial Tables

- Comparing the old and new binomial:
  - Below about 100 samples, it takes slightly fewer exceedances
     / samples to make a listing.
    - Example, if 26 samples:
      - a. Old binomial 6 or more exceed = impaired
      - b. Proposed binomial 5 or more exceed = impaired
      - c. At a straight percentage, 5 of 26 samples = 19%
  - Above about 100 samples, it takes slightly more exceedances / samples to make a listing.
    - Example, if 140 samples:
      - a. Old binomial -20 or more exceed = impaired
      - b. Proposed binomial 21 or more exceed = impaired
      - c. At a straight percentage, 21 of 140 samples = 14%

## Listing with Less Than 10 Samples

- Once 3 exceedances, why delay listing until 10 samples have been collected?
  - Provide more information about sources, extent of contamination, critical conditions, etc.
- Once 3 exceedances, benefits of listing before 10 samples:
  - It may be many months or years before further monitoring is scheduled because 40% of the monitoring data comes from other agencies.
  - Identifying sources, extent of contamination, and critical conditions is part of doing the TMDL.

## Listing With Less Than 10 Samples

### How can we avoid unnecessary listings?

- Binomial corrects for sampling error.
- Weight-of-evidence in the IWI Rule (604.B), the TMDL statute, and surface water standards provide ways to avoid listing errors. Surface waters will not be listed if:
  - a. Anthropogenic activities are not a source,
  - b. Actions have been taken in the watershed so that uses will be attaining by the next assessment,
  - c. Newer data or more reliable data exists,
  - d. More direct measurements of designated use impacts exist.
- Extensive review process.
- The regulated community can provide monitoring data from surface waters receiving discharges.

## Less than 10 Samples

# Other States Minimum Sample Size and Exceedances for Listings

States	AL	CO	KS	NH	NM	NC	OR	SC	TX	VA	WY
Minimum Samples	5	10 (4)	12 (5)	10 (2)	7 (2)	10	5	2	10	12 (2)	2
Minimum Exceed	5	4	3	2	2	2	2	2	3	2	2

Note that minimum sample size is reduced in 5 states when minimum exceedances have occurred (numbers in paranthesis apply).

## Listing With Less than 10 Samples

The probability of there being >10% exceedance rate, when you have 3 exceedances in 3-9 samples is between 99 – 100% probability!!!

(using the BINOMDIST function in Excel)

This suggests that ADEQ should make a listing with 3 exceedances and fewer than 10 samples, unless there is substantial evidence that a listing would be in error.

## Revised Binomial Approach

#### Recommendation:

Minimum sample size is 10, unless sufficient exceedances have already occurred (3 or more).

- Exceedances during at least 3 monitoring events.
- Use Table 2 if 10 or more samples.

Listings would occur when 95% confidence that more than 10% of the samples are exceeding standards and 3 or more monitoring events had exceedances.

#### Determining attainment:

- 0 exceedances in at least 3 monitoring events (seasonal distribution, core parameters, etc.) or
- If any exceedances:
  - For past assessment, fewer exceedances than required to be placed on the planning list (in the IWIR).
    - Based on 80% confidence that >10% exceedances were occurring.
  - Proposing using a binomial based on attaining at:
     Based on 85% confidence that 10% of fewer exceedances are occurring.

#### Determining attainment:

- Increase confidence level from 80% to 85%, to provide more accurate assessments.
- Calculated using CRITBINOM (sample size, 0.10, 0.85) which returns the maximum number of exceedances for a sample size, at a 10% exceedance rate with an 85% confidence level.

- Using Table 3
  - If 3 exceedances, attaining if 14 or more samples.
  - If 23 samples, attaining if only 4 or fewer exceed.

- Table 4 combines Table 2 and 3
  - A use can be assessed as impaired, attaining, or inconclusive.
  - Inconclusive if insufficient samples to determine if impaired or attaining.
  - Table 5 combines table 2 and 3
  - Example: If 20 samples

5 or more exceed = impaired

3 or fewer exceed = attaining

4 exceed = inconclusive

### Binomial for Delisting

Table 5 – Delisting <u>after</u> listed as impaired.

- The TMDL statute requires that the criteria for removing a water from the 303(d) List (delisting) cannot be more stringent than the listing criteria.
- In the past ADEQ delisted if sufficient samples to assess as "attaining."
  - If enough samples to show that it did not need to be on the Planning List (IWIR)
  - However, according to the FSU report, delisting an impaired water requires a reverse hypothesis to listing a water, and a much larger number of samples.

## Binomial for Delisting

Table 5 - Delist <u>after</u> listed as impaired.

- Reversing the hypothesis, but using the same criteria:
  - To list, assumes that the water is attaining and listing requires the minimum number of samples to show it is impaired:
    - 95% confidence level, >10% samples are exceeding standards.
  - To delist, assumes that the water is impaired and delisting requires the minimum number of samples to show it is attaining:
    - 95% confidence level, < or = 10%
    - This would require a minimum of 27 samples with 0 exceedance.
- If 0 exceed in 10 samples, using the reverse hypothesis, the confidence level drops to 65% -- this is not statistically valid!

# Binomial for Delisting

- ADEQ would support delisting based on Table 3 instead of Table 5 where:
  - Documented actions have been taken in the watershed (e.g., source removal action, best management practice implementation),
  - At least 10 samples during the critical conditions, as defined in the TMDL, and
  - Attaining uses based on Table 3 (attaining before listing).

ADEQ is recommending this alternative to further encourage that remediation actions are <u>implemented</u> in the watershed.

## **Binomial Summary**

- Use a 10 instead of 20 sample minimum for listing.
- Minimum sample size is 10 unless sufficient samples exceed standards
  - Minimum of 3 sampling events with exceedances.
- Raise confidence level from 90% to 95% in the binomial to determine an impaired use.
- Raise confidence level from 80% to 85% in the binomial to determine attaining use.
- Delisting based on same criteria but reverse hypothesis:
  - Minimum of 27 samples with 0 exceedances
  - Or can apply attaining binomial to support delisting if documented improvements in the watershed.

